

ПРИРОДНО-МАТЕМАТИЧКИ ФАКУЛТЕТ

ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ
КАНДИДАТА МИЛОША САВИЋА

I ПОДАЦИ О КОМИСИЈИ	
1.	Датум и орган који је именовео комисију 27.2.2015. године, Наставно-научно веће Природно-математичког факултета Универзитета у Новом Саду
2.	Састав комисије са знаком имена и презимена сваког члана, звања, назива уже научне области за коју је изабран у звање, датума избора у звање и назив факултета, установе у којој је члан комисије запослен: <ul style="list-style-type: none"> - др Зоран Будимац, редовни професор, <i>Рачунарске науке и информатика</i>, 15. 6. 2004. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – председник - др Мирјана Ивановић, редовни професор, <i>Рачунарске науке и информатика</i>, 29. 04. 2002. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – ментор - др Милош Радовановић, доцент, <i>Рачунарске науке</i>, 12.12.2011., Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – члан - др Бојана Димић Сурла, доцент, <i>Информациони системи</i>, 1.7.2010., Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – члан - др Зоран Огњановић, научни саветник, <i>Математичке науке, Природно-математичке науке - математика</i>, 19.11.2008, Математички институт Српске академије наука и уметности, Београд – члан
II ПОДАЦИ О КАНДИДАТУ	
1.	Име, име једног родитеља, презиме: Милош, Младен, Савић
2.	Датум рођења, општина, држава: 12.10.1984, Сарајево, Босна и Херцеговина
3.	Назив факултета, назив студијског програма дипломских академских студија – мастер и стечени стручни назив Природно-математички факултет Универзитета у Новом Саду, Информатика – мастер академске студије, Мастер информатичар
4.	Година уписа на докторске студије и назив студијског програма докторских студија 2011, Докторске академске студије – Информатика
5.	Назив факултета, назив магистарске тезе, научна област и датум одбране: /
6.	Научна област из које је стечено академско звање магистра наука: /
III НАСЛОВ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:	
Extraction and analysis of complex networks from different domains Екстракција и анализа комплексних мрежа из различитих домена	

IV ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Навести кратак садржај са знаком броја страна, поглавља, слика, шема, графикона и сл.

Докторска дисертација је написана на енглеском језику и садржи укупно xv + 215 страна А4 формата, 6 поглавља и 1 додаток, 316 библиографских референци, 56 табела и 47 слика. Поред тога дисертација садржи сажетак и проширени извод на српском језику, кратку биографију кандидата и кључну документацијску информацију. Предмет истраживања дисертације обухвата методе за екстракцију и анализу комплексних мрежа из три домена:

- мреже екстраховане из изворног кода рачунарских програма које представљају дизајн софтверских система,
- мреже екстраховане из онтологија семантичког *web*-а које описују структуру дељеног знања погодног за виšekратну употребу, и
- мреже екстраховане из библиографских записа које описују сарадњу истраживача.

Дисертација је структурирана на следећи начин:

Chapter 1: Introduction (Поглавље 1: Увод)

Chapter 2: Theoretical background (Поглавље 2: Теоријске основе)

Chapter 3: Software network (Поглавље 3: Софтверске мреже)

Chapter 4: Ontology networks (Поглавље 4: Онтолошке мреже)

Chapter 5: Co-authorship networks (Поглавље 5: Мреже сарадње истраживача)

Chapter 6: Conclusions and future work (Поглавље 6: Закључци и будући рад)

У првом поглављу дисертације се даје преглед феномена из области комплексних мрежа, а потом се уводе мреже из домена које су предмет дисертације. Друго поглавље описује теоријске основе на којима је базирана дисертација. У сваком од наредна три поглавља (поглавља 3, 4 и 5) која су посвећена конкретним доменима обухваћених дисертацијом дају се:

- Преглед релевантних истраживачких радова у области екстракције и анализе мрежа из тог конкретног домена.
- Оригинални доприноси дисертације: предлози нових техника за екстракцију мрежа из конкретног домена; методолошки оквир за анализу мрежа заснован на комбиновању поступака и метрика из теорије комплексних мрежа и метрика из домена; те конкретне анализе мрежа које представљају реалне системе из домена, а које су формиране алатима развијеним у оквиру дисертације.
- Закључак и смернице за даљи истраживачки рад у том конкретном домену.

Последње поглавље закључује дисертацију сумирајући њене резултате и доприносе, те даје сумиран преглед студија које су планиране у будућем раду.

V ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Дисертација у целини, као и њени појединачни делови, има одговарајућу структуру и план излагања. Кандидат је добро систематизовао постојеће радове у области екстракције и анализе софтверских, онтолошких и мрежа сарадње истраживача, док је својим резултатима дао оригиналан допринос области рачунарских наука. Тиме је он у потпуности реализовао постављене циљеве дисертације.

Наслов. Наслов дисертације је јасно, прецизно формулисан и одражава садржај истраживања.

Поглавље 1, Увод. У уводном поглављу дисертације дат је преглед најважнијих феномена уочених у емпиријским студијама комплексних мрежа из различитих домена те њихова веза са парадигматским моделима теорије комплексних мрежа. Потом се дефинише предмет истраживања дисертације, те се уводе појмови софтверске, онтолошке и мреже сарадње истраживача. Истичући поља примене претходно наведених мрежа кандидат даје јасну мотивацију за предмет истраживања, те његову актуелност.

Поглавље 2, Теоријске основе. У овом поглављу изложени су основни појмови теорије графова и теорије комплексних мрежа. Дате су јасне дефиниције метрика (метрике повезаности, малог света, централности, реципрочности линкова и кластерисања) и теоријских модела (моделу случајних, *small-world* и *scale-free* мрежа) које се користе у каснијим поглављима дисертације.

Поглавље 3, Софтверске мреже. У овом поглављу дисертације се прво уводи таксономија софтверских мрежа, те веза између софтверских мрежа и постојећих софтверских метрика дизајна. Потом се описује проблематика екстракције софтверских мрежа. Даје се осврт на начин екстракције софтверских мрежа у релевантним студијама које се баве статистичком анализом софтверских система. Затим се дискутује начин екстракције софтверских мрежа у постојећим језички независним алатима и окружењима за реверзни инжењеринг софтверских система. Наредна подсекција овог поглавља презентује први оригинални допринос дисертације: *SNEIPL* – проширив, језички независан приступ екстракцији софтверских мрежа који је базиран на језички независној *enriched Concrete Syntax Tree (eCST)* репрезентацији изворног кода уведеној у радовима аутора Ракић и Будимац. Даје се преглед приступа, приказ контролисаниог експеримента екстракције мрежа, приказ екстрахованих мрежа из реалних софтверских система писаних у различитим програмским језицима, те компаративна анализа са другим алатима. Следећи део овог поглавља презентује други оригинални допринос дисертације – идеју да се метрике за оцену техника кластерисања графа искористе као метрике кохезивности софтверских ентитета. Ова идеја се разматра користећи теоријски оквир за анализу софтверских метрика кохезивности који су увели *Briand* и коаутори. Претпоследња секција овог поглавља је посвећена анализи софтверских мрежа. Кандидат на јасан начин даје преглед релевантних резултата претходних истраживања из области. Потом поставља методолошки оквир за анализу мрежа које су екстраховане користећи *SNEIPL*. У постављеном методолошком оквиру кандидат анализира структуру мрежа сарадње класа екстрахованих из 5 реалних и широко коришћених Јава софтверских система отвореног кода и то:

- карактеристике слабо повезаних компоненти како би се разумео степен свеукупне кохезивности анализираних система,
- карактеристике јако повезаних компоненти како би се дубље разумео феномен цикличних зависности између ентитета који се сматра лошом одликом дизајна софтверских система,
- дистрибуције степени софтверских ентитета где се преиспитује доминантно схватање о *scale-free* структури софтверских мрежа,
- карактеристике јако повезаних класа да би се дубље разумео „*high coupling*” феномен који се такође сматра лошом одликом дизајна софтверских система.

На основу резултата анализе изводе се закључци о комплексности и квалитету дизајна анализираних софтверских система, те се даје објашњење „*high coupling*” феномена у смислу његовог узрока, те потенцијалних негативних аспеката.

Поглавље 4, Онтолошке мреже. У овом поглављу дисертације кандидат прво дискутује утицај онтолошких нормализација на предефинисану модуларизацију онтолошких описа. На бази те дискусије даје дефиницију онтолошког графа и онтолошких мрежа које се на основу њега формирају. Слично претходном поглављу дат је преглед онтолошких метрика и њихова веза са онтолошким мрежама. Као први допринос дисертације у овом сегменту кандидат претходно дискутовану идеју о употреби метрика за оцену квалитета кластерисања графа транслира у домен онтологија. Потом се дискутује проблематика екстракције онтолошких мрежа. Као оригиналан допринос тезе презентује се *ONGRAM* - екстрактор онтолошких мрежа чији су чворови обогаћени хибридном скупом метрика. *ONGRAM* је, слично *SNEIPL*-у, базиран на језички независној *eCST* репрезентацији. Истакнута је предност *eCST* репрезентације онтологија која је искоришћена како би се дефинисале нове и адаптирале постојеће софтверске метрике за евалуацију онтолошких описа. У следећој секцији поглавља даје се преглед релевантних истраживања које се баве проблематиком анализе онтолошких мрежа. Као оригиналан допринос тезе презентују се резултати анализе онтолошких мрежа које на различитим нивоима апстракције представљају *NASA*-ину модуларизовану онтологију која дефинише терминологију у домену геолошких наука. На основу резултата анализе изводе се закључци о квалитету модуларизације испитиваног онтолошког система.

Поглавље 5, Мреже сарадње истраживача. Ово поглавље дисертације је посвећено проблематици екстракције и анализе мрежа сарадње истраживача. У првој секцији поглавља се формално дефинише појам мреже сарадње истраживача. Потом се дискутује проблем екстракције мрежа сарадње истраживача. Презентују се три класе приступа идентификацији аутора у скупу библиографских записа (приступу базирани на иницијалима, хеуристички приступи и приступи

базирани на машинском учењу). Потом се елаборира њихова учесталост у студијама које се баве анализом мрежа сарадње истраживача. Специјални осврт је дат на технике идентификације аутора у масивним библиографским базама података. Следећи део поглавља нуди систематичан преглед феномена који су уочени приликом анализа мрежа сарадње истраживача у разним научним дисциплинама. Детаљан преглед студија које се баве анализом мрежа сарадње истраживача дат је за област математике, јер је та област најрелевантнија за студију случаја која је оригиналан допринос дисертације. У последњем сегменту овог поглавља се презентује студија случаја – *eLib* мрежа која описује сарадњу истраживача који су своје радове публиковали у српским математичким часописима. Радове из ових часописа индексира електронска библиотека Математичког института Српске академије наука и уметности. По нашем најбољем знању анализа мреже сарадње коју доминантно чине српски (југословенски) математичари је прва студија овакве врсте уопште, те је као таква вредан допринос разумевању социјалне структуре српске математичке баштине будући да доступни записи покривају веома широк временски период (од 1932 до данашњег дана). На основу прелиминарне анализе библиографских записа идентификују се проблеми релевантни за екстракцију *eLib* мреже. Предлаже се адекватан, полу-аутоматски приступ који укључује анализу имена засновану на употреби мера сличности стрингова. У наставку поглавља се испитују карактеристике динамике броја публикација и броја заступљених аутора на годишњем нивоу, основне карактеристике колаборативности и продуктивности аутора, структура мреже сарадње и њена еволуција. На основу резултата анализе кандидат изводи закључке о обрасцима и дугорочним трендовима истраживачке сарадње *eLib* заједнице аутора.

Поглавље 5, Закључци и будући рад. У последњем поглављу дисертације кандидат сумира оригиналне доприносе дисертације. Такође се дају смернице за будући рад.

VI СПИСАК НАУЧНИХ И СТРУЧНИХ РАДОВА КОЈИ СУ ОБЈАВЉЕНИ ИЛИ ПРИХВАЋЕНИ ЗА ОБЈАВЉИВАЊЕ НА ОСНОВУ РЕЗУЛТАТА ИСТРАЖИВАЊА У ОКВИРУ РАДА НА ДОКТОРСКОЈ ДИСЕРТАЦИЈИ

Таксативно навести називе радова, где и када су објављени. Прво навести најмање један рад објављен или прихваћен за објављивање у часопису са ISI листе односно са листе министарства надлежног за науку када су у питању друштвено-хуманистичке науке или радове који могу заменити овај услов до 01. јануара 2012. године. У случају радова прихваћених за објављивање, таксативно навести називе радова, где и када ће бити објављени и приложити потврду о томе.

M21 – Рад у врхунском међународном часопису

M. Savić, M. Ivanović, M. Radovanović, Z. Ognjanović, A. Pejović and T. Jakšić-Krüger. The structure and evolution of scientific collaboration in Serbian mathematical journals. *Scientometrics* 101(3): 1805-1830, 2014, DOI: 10.1007/s11192-014-1295-6.

M. Savić, G. Rakić, Z. Budimac and M. Ivanović. A language-independent approach to the extraction of dependencies between source code entities. *Information and Software Technology* 56(10): 1268-1288, 2014, DOI: 10.1016/j.infsof.2014.04.011.

M23 – Рад у међународном часопису

M. Savić, M. Ivanović, and M. Radovanović. Characteristics of class collaboration networks in large Java software projects. *Information Technology and Control* 40(1): 48-58, 2011.

M33 – Рад саопштен на скупу међународног значаја, штампано у целини

M. Savić, G. Rakić, Z. Budimac and M. Ivanović. Extractor of software networks from enriched concrete syntax trees. In *Proceedings of SCLIT'12, 2nd Symposium on Computer Languages, Implementations, and Tools, held within ICNAAM'12, AIP Conference Proceedings 1479*, pages 486-489, Kos, Greece, 2012, DOI: 10.1063/1.4756172.

M. Savić, Z. Budimac, G. Rakić, M. Ivanović and M. Heričko. SSQSA ontology metrics front-end. In *Proceedings of SQAMIA'13, 2nd Workshop on Software Quality Analysis, Monitoring, Improvement, and Applications, CEUR Workshop Proceedings 1053*, pages 95-101, Novi Sad, Serbia, 2013.

M. Savić and M. Ivanović. Graph clustering evaluation metrics as software metrics. In *Proceedings of SQAMIA'14, 3rd Workshop on Software Quality Analysis, Monitoring, Improvement, and Applications, CEUR Workshop Proceedings 1266*, pages 81-89, Lovran, Croatia, 2014.

M. Savić, M. Ivanović, M. Radovanović, Z. Ognjanović, A. Pejović and T. Jakšić-Krüger. Exploratory analysis of communities in co-authorship networks: a case study. In *Proceedings of ICT Innovations 2014: World of Data, Advances in Intelligent Systems and Computing 311*, pages 55-64, Ohrid, FYR Macedonia, 2014, DOI: 10.1007/978-3-319-09879-1_6.

M. Savić, M. Radovanović and M. Ivanović. Community detection and analysis of community evolution in Apache Ant class collaboration networks. In *Proceedings of BCI'12, 5th Balkan Conference in Informatics, ACM International Conference Proceedings Series 641*, pages 229-234, Novi Sad, Serbia, 2012.

Z. Budimac, G. Rakić and **M. Savić**. SSQSA architecture. In *Proceedings of BCI'12, 5th Balkan Conference in Informatics, ACM International Conference Proceedings Series 641*, pages 287-290, Novi Sad, Serbia, 2012.

VII ZAKЉUČICI OДНОСНО РЕЗУЉТАТИ ИСТРАЖИВАЊА

У погледу екстракције софтверских мрежа предложен је *SNEIPL* - нови, проширив, језички независан приступ базиран на језички независној *eCST* репрезентацији изворног кода. *SNEIPL* је реализован као један од задњих делова (енгл. *back-end*) *SSQSA (Set of Software Quality Static Analyzers)* окружења кога су развили аутори Ракић и Будимац. У контролисаном експерименту је показано да *SNEIPL* екстрахује изоморфне софтверске мреже из структурно и семантички еквивалентних програма написаних у различитим програмским језицима. Потом је приступ демонстриран екстракцијом мрежа које репрезентују реалне софтверске системе писане у различитим програмским језицима (*Java, Modula-2, Delphi*) који припадају различитим програмским парадигмама. Коректност и комплетност приступа је истраживана поређењем мрежа сарадње класа екстрахованих из 10 реалних софтвера писаних у Јави са одговарајућим мрежама екстрахованим употребом друга два алата. Наиме, у упоредној анализи су коришћени:

- *DependencyFinder*, језички зависан алат који мрежне репрезентације Јава софтверских система формира на основу њиховог бајт кода, и
- *Doxygen*, језички независан екстрактор базиран на фази плитком парсирању. Додатно се наводе примери коришћења овог алата у експерименталним истраживањима у софтверском инжењерству како би се показала његова распрострањеност у академској заједници.

Упоредна анализа је показала да су мреже екстраховане *SNEIPL*-ом сличне онима које су екстраховане коришћењем *DependencyFinder* (преко 90% сличности у свим студијама случаја), те да *SNEIPL* даје далеко прецизније резултате него *Doxygen*. Како се на основу софтверских мрежа рачунају софтверске метрике које квантификују комплексност и квалитет дизајна софтверских система, то *SNEIPL* омогућава и језички-независно рачунање одређене класе софтверских метрика. Анализом постојећих, широко коришћених, језички независних алата и окружења за реверзни инжењеринг софтверских система (*Rigi, Moose, Gupro, Bauhaus*) утврђено је да исти омогућавају језички независну репрезентацију екстрахованих факата (мрежних модела изворног кода), али не и језички независну екстракцију факата. Наиме поменути алата су тако реализовани да за сваки подржани програмски језик постоји засебан екстрактор факата или је екстракција факата парцијално језички независна у смислу да се софтверске мреже екстрахују из језички-независне репрезентације ниског нивоа само за одређен подкуп подржаних језика. Стога *SNEIPL*, у односу на претходно поменута окружења и алата, омогућава језички-независан приступ формирању базе факата које се користе у реверзном инжењерингу софтверских система.

„*Low coupling-high cohesion*” је један од базичних принципа софтверског и онтолошког инжењерства чија примена води квалитетној и одрживој модуларизацији система. Кандидат

правилно уочава да дословна примена тог принципа мора резултовати кластер структуром софтверских и онтолошких мрежа. Кластер структура мреже подразумева да у мрежи постоје кохезивне групе чворова које су гушће повезане између себе него са остатком мреже. Ако је неки софтверски/онтолошки ентитет јако кохезиван тада се ентитети које он дефинише појављују као кластер у одговарајућој софтверској/онтолошкој мрежи. Стога се предлажу метрике за оцену техника кластерисања графа (*GCE* метрике) као софтверске и онтолошке метрике кохезивности. За разлику од стандардних метрика кохезивности које се користе у софтверском/онтолошком инжењерству, метрике кластерисања графа не игноришу референце на екстерне ентитете (ентитете ван модула), те самим тим не оцењују кохезивност ентитета у изолацији од остатка система. Теоријском анализом *GCE* метрика као софтверских метрика у оквиру који су предложили *Briand* и коаутори показано је да разматране *GCE* метрике задовољавају две најважније особине добро дефинисаних кохезивних метрика: монотоност и особина спајања. Ове метрике међутим поседују извесна ограничења којих морамо бити свесни приликом њихове употребе. Емпиријска евалуација у домену модуларизованих онтологија је показала да не постоји корелација између *GCE* метрика и метрика кохезивности које се ослањају само на интерне референце (референце унутар модула) што самим тим указује на корисност *GCE* метрика као комплементарних метрика стандардним метрикама кохезивности.

У погледу екстракције онтолошких мрежа предложен је *ONGRAM* – нови, проширив, језички независан приступ који је, слично *SNEIPL*-у, базиран на *eCST* репрезентацији и реализован као један од задњих делова *SSQSA* окружења. *ONGRAM* омогућује проширивост на различите језике за репрезентацију знања: проширивост *SNEIPL*-а и *ONGRAM*-а је директна последица проширивости *SSQSA* окружења. Поврх тога, *eCST* репрезентација онтолошких описа је омогућила дефинисање нових и адаптацију постојећих софтверских метрика унутрашње комплексности. Наиме, показано је како се скуп *Halstead*-ових метрика и *Henry-Kafura* комплексност могу адаптирати за евалуацију онтологија, те је предложена нова метрика комплексности израза (енгл. *expression complexity*) која квантификује комплексност *OWL* аксиома. На основу екстрахованих онтолошких мрежа *ONGRAM* такође рачуна:

- претходно уведене онтолошке метрике дизајна,
- метрике независне од домена које се примењују у анализи социјалних мрежа (метрике централности),
- *GCE* метрике, и
- хибридне метрике које комбинују метрике унутрашње комплексности и метрике дизајна.

Стога, *ONGRAM* формира онтолошке мреже чији су чворови проширени богатим, хибридним скупом метрика.

У дисертацији је предложен полу-аутоматски приступ екстракцији мрежа сарадње истраживача погодан за ретке и фрагментисане мреже базиран на хеуристикама за идентификацију именских хомонима и синонима. Именски хомоними се утврђују на основу структуре его мреже за аутора. Наиме, уколико је аутор артикулациони чвор у својој его мрежи тада се ручно проверава да ли аутор заправо представља више истраживача који имају исто име. Именски синоними се одређују у две фазе на основу постојећих мера сличности стрингова и класификације имена у две категорије. Прву категорију имена чине она имена где је презиме дато у пуном облику, а лично име скраћено на иницијале. У другу категорију имена спадају она имена где су и лично име и презиме дати у пуном облику. У првој фази се мере сличности стрингова примењују без ограничења (без примене неке блокирајуће функције). У другој фази се примењује блокирајућа функција која узрокује да се метрике сличности рачунају само за парове аутора из исте повезане компоненте у мрежи сарадње. Приступ је примењиван у екстракцији мреже сарадње истраживача из библиографских записа електронске библиотеке Математичког института Српске академије науке и уметности (*eLib*). Мрежа је потом анализирана како би се идентификовали обрасци и дугорочни трендови сарадње аутора који су своје радове публиковали у српским математичким часописима.

Користећи алате за екстракцију мрежа који су развијени у оквиру дисертације формиран је експериментални скуп мрежа које су потом анализирани. Експериментални скуп мрежа се састоји од:

- 5 мрежа сарадње класа софтверских система отвореног кода реализованих у програмском

језику Јава: *Tomcat, Lucene, Ant, Xerces, JFreeChart,*

- 3 онтолошке мреже које на различитим нивоима апстракције представљају модуларизовану онтологију која дефинише терминологију у домену геолошких наука,
- мрежу сарадње истраживача екстраховану из *eLib* дигиталне библиотеке.

Јединствен карактер анализа спроведених у дисертацији је да комбинују поступке и метрике из теорије комплексних мрежа са метрикама из конкретног домена (софтверске и онтолошке метрике, метрике продуктивности и дугорочне заступљености истраживача). Предложен је статистички тест који пореди две групе чворова у мрежи при чему се групе формирају на бази неког тополошког критеријума, а сваки чвор је описан неким метричким вектором. Тест се заснива на секвенцијалној (секвенцијалној по елементима метричког вектора) употреби непараметарског *Mann-Whitney U* теста и вероватноћама супериорности. На основу теста се идентификују оне метрике за које не постоје статистички значајне разлике у метричким вредностима за поређене групе чворова, метрике где постоје статистички значајне разлике, и метрике где постоје драстичне статистички значајне разлике (степен „драстичности“ статистички значајних разлика се утврђује на основу добијених вероватноћа супериорности). Показано је да мреже из експерименталног корпуса поседују *Watts-Strogatz* особину малог света која се читава релативно кратким дистанцама између чворова у мрежи и високим степеном локалног кластерисања. Друге особине (слабо) повезаних компоненти као што су степен њихове јаке повезаности, обрасци асоративности, дистрибуције степени чвора, те карактеристике јако повезаних чворова варирају кроз домене, системе и степене апстракције рефлектујући специфичности индивидуалних система.

Анализа јако повезаних компоненти присутних у испитиваним софтверским мрежама је показала да оне теже да се згушњавају са величином компоненте – просечан степен чвора унутар компоненте расте са величином компоненте. Феномен згушњавања се може описати степеним законом (енгл. *power-law*) чији се експоненти могу искористити као индикатори квалитета дизајна софтверског система. Код софтверских и онтолошких мрежа је такође примећено да постоји јак дисбаланс између улазног и излазног степена јако повезаних чворова (чворова са великим укупним степеном, при чему је укупни степен збир улазног и излазног степена). Овај резултат имплицира да су јако повезани ентитет софтверских и онтолошких система узроковани доминатно њиховом поновним искоришћењем (енгл. *internal reuse*). Како се јако повезани ентитети сматрају негативном појавом у домену софтверског и онтолошког инжењерства, то наведени резултат надаље повлачи да присутност високог степена везивања може сугерисати само негативне аспекте поновног искоришћења ентитета, не и негативне аспекте агрегације ентитета. Применом теста (базираног на метрикама) поређења група чворова утврђене су детаљне карактеристике јако повезаних компоненти и јако повезаних чворова испитиваних софтверских и онтолошких мрежа. Наиме, показано је да јако повезани чворови софтверских и онтолошких система (ОО класе и онтолошки модули) имају већи волумен и глобалну централност у односу на остатак ентитета, те да испитиван онтолошки систем има јако повезано језгро (енгл. *strongly connected core*) које сачињавају најважнији и најискоришћенији онтолошки модули.

За разлику од испитиваних софтверских и онтолошких мрежа које имају гигантску слабо повезану компоненту, анализа компоненти *eLib* мреже је открила тополошку разноврсност у структури мреже: мрежа се састоји из релативно великог броја повезаних компоненти које су у складу са степеним законом, нема гигантску повезану компоненту, већину компоненти чине изоловани аутори и мале тривијалне компоненте, али такође постоји и мали број релативно великих, нетривијалних компоненти повезаних аутора. Применом техника за детекцију заједница је показано да највеће компоненте у мрежи имају јасну структуру заједница, односно да се у мрежи могу уочити јако кохезивне групе аутора. Испитивањем структуре подмрежа *eLib* мреже утврђене су карактеристике заједница које одговарају појединачним часописима које *eLib* индексира. Анализа еволуције мреже је показала да постоји шест карактеристичних периода у еволуцији мреже које се разликују по интензитету и типу колаборативног понашања аутора. Анализом метрика на нивоу аутора је показано да је угњежденост аутора бољи показатељ продуктивности и заступљености у *eLib* часописима него број коаутора.

VIII ОЦЕНА НАЧИНА ПРИКАЗА И ТУМАЧЕЊА РЕЗУЛТАТА ИСТРАЖИВАЊА

Експлицитно навести позитивну или негативну оцену начина приказа и тумачења резултата истраживања.

Комисија позитивно оцењује начин приказа и тумачња резултата. Оригинални резултати дисертације приказани су систематично и јасно, уз употребу садржајних табела и графикана, тумаче се коректно, и дискутују у склопу претходних релевантних истраживања.

IX КОНАЧНА ОЦЕНА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Експлицитно навести да ли дисертација јесте или није написана у складу са наведеним образложењем, као и да ли она садржи или не садржи све битне елементе. Дати јасне, прецизне и концизне одговоре на 3. и 4. питање:

1. Да ли је дисертација написана у складу са образложењем наведеним у пријави теме

Увидом у образложење наведено у пријави теме може се утврдити да је дисертација написана у складу са планом истраживања наведеним у пријави теме и да добијени резултати одговарају циљевима постављеним у пријави теме.

2. Да ли дисертација садржи све битне елементе

Дисертација садржи све битне елементе научно-истраживачког рада: неопходну теоријску основу, приказ релевантних истраживања из области, и оригиналне методолошке и емпиријске доприносе који унапређују постојећа знања и схватања у домену екстракције и анализе софтверских, онтолошких и мрежа сарадње истраживача. Обиман списак библиографских референци (316 од чега је кандидат коаутор у 9 референци) садржи релевантне радове и сведочи да је кандидат добро упознао област истраживања.

3. По чему је дисертација оригиналан допринос науци

Оригинални доприноси дисертације су:

1. Дизајн и имплементација проширеног, језички-независног приступа екстракцији софтверских мрежа базираног на *enriched Concrete Syntax tree* репрезентацији изворног кода.
2. Дизајн и имплементација проширеног, језички-независног приступа екстракцији онтолошких мрежа чији су чворови обогаћени хибридном скупом метрика, базираног на *enriched Concrete Syntax tree* репрезентацији онтолошких описа.
3. Увођење нових онтолошких и адаптација постојећих софтверских метрика унутрашње комплексности за евалуацију онтолошких описа, увођење метрика за оцену техника кластерисања графа као софтверских и онтолошких метрика.
4. Увођење статистичког теста за поређење група чворова у мрежи, при чему је сваки чвор окарактерисан неким метричким вектором, а групе се формирају на бази неког тополошког критеријума.
5. Дизајн и имплементација полу-аутоматског приступа екстракцији мрежа сарадње истраживача погодног за ретке и фрагментисане мреже базираног на хеуристикама за идентификацију именских хомонима и синонима.
6. Анализа корпуса мрежа екстрахованих применом алата развијених у оквиру дисертације. У односу на претходне анализе мрежа из домена софтверских, онтолошких и социјалних система, анализе представљене у дисертацији комбинују поступке и метрике развијене у оквиру теорије комплексних мрежа са метрикама из конкретних домена (софтверским, онтолошким и метрикама продуктивности и дугорочне заступљености истраживача), те на тај начин дају дубљу карактеризацију структуре и еволуције анализираних система.

4. Недостаци дисертације и њихов утицај на резултат истраживања

Дисертација нема недостатака.

X ПРЕДЛОГ:

На основу укупне оцене дисертације, комисија предлаже: да се докторска дисертација под називом **„Extraction and analysis of complex networks from different domains” (Екстракција и анализа комплексних мрежа из различитих домена)** кандидата **Милоша Савића** прихвати, а кандидату одобри јавна одбрана дисертације.

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

др Зоран Будимац, редовни професор, председник

др Мирјана Ивановић, редовни професор, ментор

др Милош Радовановић, доцент, члан

др Бојана Димић Сурла, доцент, члан

др Зоран Огњановић, научни саветник, члан

НАПОМЕНА: Члан комисије који не жели да потпише извештај јер се не слаже са мишљењем већине чланова комисије, дужан је да унесе у извештај образложење односно разлоге због којих не жели да потпише извештај.